

## RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/782,401A  
Source: IFWB  
Date Processed by STIC: 1-6-05

# ***ENTERED***



IFWO

## RAW SEQUENCE LISTING

DATE: 01/06/2005

PATENT APPLICATION: US/10/782,401A

TIME: 16:16:18

Input Set : E:\ARCD390US.APP.txt

Output Set: N:\CRF4\01062005\J782401A.raw

```

3 <110> APPLICANT: ASHTON-RICKARDT, PHILIP
5 <120> TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE INHIBITION OF
6   CATHEPSINS
8 <130> FILE REFERENCE: ARCD:390US
10 <140> CURRENT APPLICATION NUMBER: 10/782,401A
11 <141> CURRENT FILING DATE: 2004-02-19
13 <150> PRIOR APPLICATION NUMBER: 60/448,285
14 <151> PRIOR FILING DATE: 2003-02-19
16 <160> NUMBER OF SEQ ID NOS: 25
18 <170> SOFTWARE: PatentIn Ver. 2.1
20 <210> SEQ ID NO: 1
21 <211> LENGTH: 1323
22 <212> TYPE: DNA
23 <213> ORGANISM: Mus musculus
25 <400> SEQUENCE: 1
26 atggctggtg tctcccctgc tgtctttggc tgcccagatg tcaccctggg aaggaacact 60
27 gcagtccgtg aagtacaaga aaatgtcaca tcagtggaca gtttaacact ggtctccagc 120
28 aacactgact ttgccttcag cctctacagg aagctggttt tgaagaatcc agatgaaaaat 180
29 gttgtcttct ccccatcag catctgcact gccttggccc tcctgtccct gggagcaaaag 240
30 agcaacaccc tgaaggaaat cctagaaggt ctcaagttca acctcacaga gaccctgaa 300
31 ccagacatcc accagggtt taggtacttg ctagacctt taagtcagcc agggaaccag 360
32 gtacagatca gcacaggcag tgcctgttt attgaaaagc acctacagat cctggcagag 420
33 ttcaaggaga aggcaagggc tctgtaccag gctgaggcct tcacagcaga tttccagcaa 480
34 cctctcaagg ccacaaagct catcaatgac tatgtgagca atcacaccca ggggaagatc 540
35 aaggaactca tctcaggcct gaaagagagc acgttgatgg tgctggtgaa ctacatctac 600
36 tttaaaggca aatggaagaa cccctttgac ccgaatgata catttaagtc cgagttctac 660
37 ttggatgaga agaggctgt gattgtgtcc atgatgaaaa ctggttacct gacgacaccc 720
38 tacttccggg atgaggagct gtcctgcact gtggtggagc tgaagtacac aggaaatgcc 780
39 agtgccatgt tcacctccc tgaccagggc aggatgcagc aggtggaagc aagcttacia 840
40 ccagagaccc tgagggaagt gaagaactct ctgaagccca ggatgataca tgagctccgc 900
41 ctgccaagt tctccatctc caccgactat agcctggagc acatccttcc tgagttgggc 960
42 atcaggaag tcttctccac acacgctgac ctgtctgcaa tcacaggaac taaggatctg 1020
43 agagtctctc aggtggtcca caaggctgtg ctggatgtgg ctgagaaagg cacagaggct 1080
44 gctgctgcca caggaatggc aggtgtcgga tgttgtgcag tttttgactt tctggaaata 1140
45 tttttcaaca ggccattcct gatgattatc tctgacacaa aagctcacat tgccctcttt 1200
46 atggcaaaaag ttacaaatcc agagagatct acgaacttcc caaatggtga ggggtgcttct 1260
47 tccaacggc tggaatctaa gcgtttatgt tttggtgatc cgctatgcct gattggacag 1320
48 tga
51 <210> SEQ ID NO: 2
52 <211> LENGTH: 440
53 <212> TYPE: PRT
54 <213> ORGANISM: Mus musculus
56 <400> SEQUENCE: 2

```

## RAW SEQUENCE LISTING

DATE: 01/06/2005

PATENT APPLICATION: US/10/782,401A

TIME: 16:16:18

Input Set : E:\ARCD390US.APP.txt

Output Set: N:\CRF4\01062005\J782401A.raw

```

57 Met Ala Gly Val Ser Pro Ala Val Phe Gly Cys Pro Asp Val Thr Leu
58   1           5           10           15
60 Gly Arg Asn Thr Ala Val Arg Glu Val Gln Glu Asn Val Thr Ser Val
61           20           25           30
63 Asp Ser Leu Thr Leu Val Ser Ser Asn Thr Asp Phe Ala Phe Ser Leu
64           35           40           45
66 Tyr Arg Lys Leu Val Leu Lys Asn Pro Asp Glu Asn Val Val Phe Ser
67           50           55           60
69 Pro Phe Ser Ile Cys Thr Ala Leu Ala Leu Leu Ser Leu Gly Ala Lys
70   65           70           75           80
72 Ser Asn Thr Leu Lys Glu Ile Leu Glu Gly Leu Lys Phe Asn Leu Thr
73           85           90           95
75 Glu Thr Pro Glu Pro Asp Ile His Gln Gly Phe Arg Tyr Leu Leu Asp
76           100          105          110
78 Leu Leu Ser Gln Pro Gly Asn Gln Val Gln Ile Ser Thr Gly Ser Ala
79           115          120          125
81 Leu Phe Ile Glu Lys His Leu Gln Ile Leu Ala Glu Phe Lys Glu Lys
82           130          135          140
84 Ala Arg Ala Leu Tyr Gln Ala Glu Ala Phe Thr Ala Asp Phe Gln Gln
85 145           150           155           160
87 Pro Leu Lys Ala Thr Lys Leu Ile Asn Asp Tyr Val Ser Asn His Thr
88           165          170          175
90 Gln Gly Lys Ile Lys Glu Leu Ile Ser Gly Leu Lys Glu Ser Thr Leu
91           180          185          190
93 Met Val Leu Val Asn Tyr Ile Tyr Phe Lys Gly Lys Trp Lys Asn Pro
94           195          200          205
96 Phe Asp Pro Asn Asp Thr Phe Lys Ser Glu Phe Tyr Leu Asp Glu Lys
97           210          215          220
99 Arg Ser Val Ile Val Ser Met Met Lys Thr Gly Tyr Leu Thr Thr Pro
100 225           230           235           240
102 Tyr Phe Arg Asp Glu Glu Leu Ser Cys Thr Val Val Glu Leu Lys Tyr
103           245           250           255
105 Thr Gly Asn Ala Ser Ala Met Phe Ile Leu Pro Asp Gln Gly Arg Met
106           260           265           270
108 Gln Gln Val Glu Ala Ser Leu Gln Pro Glu Thr Leu Arg Lys Trp Lys
109           275           280           285
111 Asn Ser Leu Lys Pro Arg Met Ile His Glu Leu Arg Leu Pro Lys Phe
112           290           295           300
114 Ser Ile Ser Thr Asp Tyr Ser Leu Glu His Ile Leu Pro Glu Leu Gly
115 305           310           315           320
117 Ile Arg Glu Val Phe Ser Thr His Ala Asp Leu Ser Ala Ile Thr Gly
118           325           330           335
120 Thr Lys Asp Leu Arg Val Ser Gln Val Val His Lys Ala Val Leu Asp
121           340           345           350
123 Val Ala Glu Lys Gly Thr Glu Ala Ala Ala Thr Gly Met Ala Gly
124           355           360           365
126 Val Gly Cys Cys Ala Val Phe Asp Phe Leu Glu Ile Phe Phe Asn Arg
127           370           375           380
129 Pro Phe Leu Met Ile Ile Ser Asp Thr Lys Ala His Ile Ala Leu Phe

```

## RAW SEQUENCE LISTING

DATE: 01/06/2005

PATENT APPLICATION: US/10/782,401A

TIME: 16:16:18

Input Set : E:\ARCD390US.APP.txt

Output Set: N:\CRF4\01062005\J782401A.raw

```

130 385          390          395          400
132 Met Ala Lys Val Thr Asn Pro Glu Arg Ser Thr Asn Phe Pro Asn Gly
133          405          410          415
135 Glu Gly Ala Ser Ser Gln Arg Leu Glu Ser Lys Arg Leu Cys Phe Gly
136          420          425          430
138 Asp Pro Leu Cys Leu Ile Gly Gln
139          435          440
142 <210> SEQ ID NO: 3
143 <211> LENGTH: 379
144 <212> TYPE: PRT
145 <213> ORGANISM: Homo sapiens
147 <400> SEQUENCE: 3
148 Met Glu Gln Leu Ser Ser Ala Asn Thr Arg Phe Ala Leu Asp Leu Phe
149 1          5          10          15
151 Leu Ala Leu Ser Glu Asn Asn Pro Ala Gly Asn Ile Phe Ile Ser Pro
152          20          25          30
154 Phe Ser Ile Ser Ser Ala Met Ala Met Val Phe Leu Gly Thr Arg Gly
155          35          40          45
157 Asn Thr Ala Ala Gln Leu Ser Lys Thr Phe His Phe Asn Thr Val Glu
158          50          55          60
160 Glu Val His Ser Arg Phe Gln Ser Leu Asn Ala Asp Ile Asn Lys Arg
161 65          70          75          80
163 Gly Ala Ser Tyr Ile Leu Lys Leu Ala Asn Arg Leu Tyr Gly Glu Lys
164          85          90          95
166 Thr Tyr Asn Phe Leu Pro Glu Phe Leu Val Ser Thr Gln Lys Thr Tyr
167          100          105          110
169 Gly Ala Asp Leu Ala Ser Val Asp Phe Gln His Ala Ser Glu Asp Ala
170          115          120          125
172 Arg Lys Thr Ile Asn Gln Trp Val Lys Gly Gln Thr Glu Gly Lys Ile
173          130          135          140
175 Pro Glu Leu Leu Ala Ser Gly Met Val Asp Asn Met Thr Lys Leu Val
176 145          150          155          160
178 Leu Val Asn Ala Ile Tyr Phe Lys Gly Asn Trp Lys Asp Lys Phe Met
179          165          170          175
181 Lys Glu Ala Thr Thr Asn Ala Pro Phe Arg Leu Asn Lys Lys Asp Arg
182          180          185          190
184 Lys Thr Val Lys Met Met Tyr Gln Lys Lys Lys Phe Ala Tyr Gly Tyr
185          195          200          205
187 Ile Glu Asp Leu Lys Cys Arg Val Leu Glu Leu Pro Tyr Gln Gly Glu
188          210          215          220
190 Glu Leu Ser Met Val Ile Leu Leu Pro Asp Asp Ile Glu Asp Glu Ser
191 225          230          235          240
193 Thr Gly Leu Lys Lys Ile Glu Glu Gln Leu Thr Leu Glu Lys Leu His
194          245          250          255
196 Glu Trp Thr Lys Pro Glu Asn Leu Asp Phe Ile Glu Val Asn Val Ser
197          260          265          270
199 Leu Pro Arg Phe Lys Leu Glu Glu Ser Tyr Thr Leu Asn Ser Asp Leu
200          275          280          285
202 Ala Arg Leu Gly Val Gln Asp Leu Phe Asn Ser Ser Lys Ala Asp Leu

```

## RAW SEQUENCE LISTING

DATE: 01/06/2005

PATENT APPLICATION: US/10/782,401A

TIME: 16:16:18

Input Set : E:\ARCD390US.APP.txt

Output Set: N:\CRF4\01062005\J782401A.raw

```

203      290      295      300
205 Ser Gly Met Ser Gly Ala Arg Asp Ile Phe Ile Ser Lys Ile Val His
206 305      310      315      320
208 Lys Ser Phe Val Glu Val Asn Glu Glu Gly Thr Glu Ala Ala Ala Ala
209      325      330      335
211 Thr Ala Gly Ile Ala Thr Phe Cys Met Leu Met Pro Glu Glu Asn Phe
212      340      345      350
214 Thr Ala Asp His Pro Phe Leu Phe Phe Ile Arg His Asn Ser Ser Gly
215      355      360      365
217 Ser Ile Leu Phe Leu Gly Arg Phe Ser Ser Pro
218      370      375
221 <210> SEQ ID NO: 4
222 <211> LENGTH: 415
223 <212> TYPE: PRT
224 <213> ORGANISM: Homo sapiens
226 <400> SEQUENCE: 4
227 Met Glu Asp Leu Cys Val Ala Asn Thr Leu Phe Ala Leu Asn Leu Phe
228 1 5 10 15
230 Lys His Leu Ala Lys Ala Ser Pro Thr Gln Asn Leu Phe Leu Ser Pro
231 20 25 30
233 Trp Ser Ile Ser Ser Thr Met Ala Met Val Tyr Met Gly Ser Arg Gly
234 35 40 45
236 Ser Thr Glu Asp Gln Met Ala Lys Val Leu Gln Phe Asn Glu Val Gly
237 50 55 60
239 Ala Asn Ala Val Thr Pro Met Thr Pro Glu Asn Phe Thr Ser Cys Gly
240 65 70 75 80
242 Phe Met Gln Gln Ile Gln Lys Gly Ser Tyr Pro Asp Ala Ile Leu Gln
243 85 90 95
245 Ala Gln Ala Ala Asp Lys Ile His Ser Ser Phe Arg Ser Leu Ser Ser
246 100 105 110
248 Ala Ile Asn Ala Ser Thr Gly Asn Tyr Leu Leu Glu Ser Val Asn Lys
249 115 120 125
251 Leu Phe Gly Glu Lys Ser Ala Ser Phe Arg Glu Glu Tyr Ile Arg Leu
252 130 135 140
254 Cys Gln Lys Tyr Tyr Ser Ser Glu Pro Gln Ala Val Asp Phe Leu Glu
255 145 150 155 160
257 Cys Ala Glu Glu Ala Arg Lys Lys Ile Asn Ser Trp Val Lys Thr Gln
258 165 170 175
260 Thr Lys Gly Lys Ile Pro Asn Leu Leu Pro Glu Gly Ser Val Asp Gly
261 180 185 190
263 Asp Thr Arg Met Val Leu Val Asn Ala Val Tyr Phe Lys Gly Lys Trp
264 195 200 205
266 Lys Thr Pro Phe Glu Lys Lys Leu Asn Gly Leu Tyr Pro Phe Arg Val
267 210 215 220
269 Asn Ser Ala Gln Arg Thr Pro Val Gln Met Met Tyr Leu Arg Glu Lys
270 225 230 235 240
272 Leu Asn Ile Gly Tyr Ile Glu Asp Leu Lys Ala Gln Ile Leu Glu Leu
273 245 250 255
275 Pro Tyr Ala Gly Asp Val Ser Met Phe Leu Leu Leu Pro Asp Glu Ile

```

## RAW SEQUENCE LISTING

DATE: 01/06/2005

PATENT APPLICATION: US/10/782,401A

TIME: 16:16:18

Input Set : E:\ARCD390US.APP.txt

Output Set: N:\CRF4\01062005\J782401A.raw

```

276          260          265          270
278 Ala Asp Val Ser Thr Gly Leu Glu Leu Leu Glu Ser Glu Ile Thr Tyr
279          275          280          285
281 Asp Lys Leu Asn Lys Trp Thr Ser Lys Asp Lys Met Ala Glu Asp Glu
282          290          295          300
284 Val Glu Val Tyr Ile Pro Gln Phe Lys Leu Glu Glu His Tyr Glu Leu
285 305          310          315          320
287 Arg Ser Ile Leu Arg Ser Met Gly Met Glu Asp Ala Phe Asn Lys Gly
288          325          330          335
290 Arg Ala Asn Phe Ser Gly Met Ser Glu Arg Asn Asp Leu Phe Leu Ser
291          340          345          350
293 Glu Val Phe His Gln Ala Met Val Asp Val Asn Glu Glu Gly Thr Glu
294          355          360          365
296 Ala Ala Ala Gly Thr Gly Gly Val Met Thr Gly Arg Thr Gly His Gly
297          370          375          380
299 Gly Pro Gln Phe Val Ala Asp His Pro Phe Leu Phe Leu Ile Met His
300 385          390          395          400
302 Lys Ile Thr Asn Cys Ile Leu Phe Phe Gly Arg Phe Ser Ser Pro
303          405          410          415
306 <210> SEQ ID NO: 5
307 <211> LENGTH: 390
308 <212> TYPE: PRT
309 <213> ORGANISM: Homo sapiens
311 <400> SEQUENCE: 5
312 Met Asn Ser Leu Ser Glu Ala Asn Thr Lys Phe Met Phe Asp Leu Phe
313 1          5          10          15
315 Gln Gln Phe Arg Lys Ser Lys Glu Asn Ile Phe Tyr Ser Pro Ile
316          20          25          30
318 Ser Ile Thr Ser Ala Leu Gly Met Val Leu Leu Gly Ala Lys Asp Asn
319          35          40          45
321 Thr Ala Gln Gln Ile Lys Lys Val Leu His Phe Asp Gln Val Thr Glu
322          50          55          60
324 Asn Thr Thr Gly Lys Ala Ala Thr Tyr His Val Asp Arg Ser Gly Asn
325 65          70          75          80
327 Val His His Gln Phe Gln Lys Leu Leu Thr Glu Phe Asn Lys Ser Thr
328          85          90          95
330 Asp Ala Tyr Glu Leu Lys Ile Ala Asn Lys Leu Phe Gly Glu Lys Thr
331          100          105          110
333 Tyr Leu Phe Leu Gln Glu Tyr Leu Asp Ala Ile Lys Lys Phe Tyr Gln
334          115          120          125
336 Thr Ser Val Glu Ser Val Asp Phe Ala Asn Ala Pro Glu Glu Ser Arg
337          130          135          140
339 Lys Lys Ile Asn Ser Trp Val Glu Ser Gln Thr Asn Glu Lys Ile Lys
340 145          150          155          160
342 Asn Leu Ile Pro Glu Gly Asn Ile Gly Ser Asn Thr Thr Leu Val Leu
343          165          170          175
345 Val Asn Ala Ile Tyr Phe Lys Gly Gln Trp Glu Lys Lys Phe Asn Lys
346          180          185          190
348 Glu Asp Thr Lys Glu Glu Lys Phe Trp Pro Asn Lys Asn Thr Tyr Lys

```

**VERIFICATION SUMMARY**

DATE: 01/06/2005

PATENT APPLICATION: US/10/782,401A

TIME: 16:16:19

Input Set : E:\ARCD390US.APP.txt

Output Set: N:\CRF4\01062005\J782401A.raw